DOCUMENT RESUME

ED 169 111 TM 008 549

AUTHOR Morris, John D.
TITLE Comparison of Predictive Accuracy on Several Types of Factor Scores.
PUB DATE Aug 78
NOTE 17p.; Paper presented at the Annual Meeting of the American Psychological Association (86th, Toronto, Ontario, Canada, August, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Statistics; Correlation; *Factor Analysis; *Multiple Regression Analysis; Predictive Measurement; Predictive Validity; *Scores; Simulation; *Weighted Scores
IDENTIFIERS *Factor Scores

ABSTRACT

Several advantages to the use of factor scores as independent variables in a multiple regression equation were found. To help select the most desirable type of factor score on which to calculate a regression equation, computer-based Monte Carlo methods were used to compare the predictive accuracy upon replication of regression of five "complete" and four "incomplete" factor score estimation methods. For several levels of multiple correlation (R2=.30, .50, and .70), and for several subject-to-variable sampling ratios (3/1, 5/1, and 10/1), prediction on incomplete factor scores showed significantly better double cross-validated prediction accuracy than on complete factor scores. Moreover, the unique unit-weighted factor score was superior among the incomplete methods. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED169111

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Comparison of Predictive Accuracy on Several Types of Factor Scores

by

John D. Morris
Georgia Southern College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John D. Morris

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Paper Presented at the Annual Meeting
of the American Psychological Association
Division 5

Toronto, Ontario, Canada

August, 1978

(Printed in U.S.A.)

IM008 549

Abstract

Several advantages to the use of factor scores as independent variables in a multiple regression equation are reported. In an effort to help the researcher select the most desirable type of factor score on which to calculate a regression equation, computer based Monte Carlo methods were used to compare the predictive accuracy upon replication of regression on five "complete" and four "incomplete" factor score estimation methods using a data set from the literature. For several levels of multiple correlation (R^2 = .30, .50, and .70), and for several subject to variable sampling ratios (3/1, 5/1, and 10/1), prediction on incomplete factor scores showed significantly ($p < .001$) better double cross-validated prediction accuracy than on complete factor scores. Moreover, the unique unit-weighted factor score was superior among the incomplete methods.

A Comparison of Regression Prediction Accuracy on Several Types of Factor Scores

The drastic deflation in a multiple correlation coefficient which often arises in applying a regression equation to a replication sample sometimes leads investigating scientists into a position of disfavor with the entire prediction process. Yet such examination of the robustness of the sample specific regression weights is absolutely mandatory for any assurance of predictive generalizability, surely a primary goal in any science. The most fundamental solution is to include a larger and more representative sample in the group upon which the predictive model is built, however, some authors (Cattell, 1966, p. 80; Guertin and Bailey, 1970, p. 196) have also suggested that factor analyzing the full-rank data and using factor scores as predictors rather than the full-rank data provides for more generalizable replication prediction.

The replication predictive accuracy of regression equations on common factor scores has been empirically examined as a function of sample size (Burket, 1964), and in relation to the degree of similarity between the replication and equation building samples (Morris, 1977; Morris & Guertin, 1977). Results have suggested superiority over full-rank data for the traditional factor scores estimated by regression under certain conditions.

In addition, regression on factor scores rather than on full-rank data has been recommended because of increased error degrees of freedom (Horst, 1965). In the case of orthogonal factor solutions, regression on factor scores has been suggested to avoid the problems associated with correlated predictor variables (Kerlinger & Pedhazur, 1973); this being especially relevant in the case of multicolinearity.

Several alternative methods for deriving weights for the linear composite factor score estimates are used and have been compared for orthogonal factor rotations on desirable internal attributes of validity (high correlation with the appropriate true factor score), univocality (low correlation with other true factor scores), and orthogonality (low factor score estimate intercorrelation). Horn (1965) classified these different methods as "complete" (some least squares procedure involved), and "incomplete" (weights derived directly from the factor structure in some intuitive manner).

Theoretical examination of complete factor score estimates has pointed out that none of the methodologies produce factor scores in an orthogonal common factor solution that posses all three of these desirable characteristics (McDonald and Burr, 1967), but that when factor scores are to be used as independent variables in a regression equation, traditional estimates by regression have some theoretical advantage over the other types of complete estimates (Tucker, 1971).

Empirical examination of the internal characteristics of complete vs. incomplete factor scores in a variety of common and component solutions (Glass and Maguire, 1966; Halperin, 1976; Horn, 1965; Horn and Miller, 1966; Moseley and Klett, 1964) seems to support the consensus that, while incomplete factor scores are reasonably highly correlated with the appropriate complete factor score, they are also often highly correlated with each other, and with inappropriate complete factor scores. Therefore, with differing degrees of fervor, most of these authors have suggested that incomplete estimation methods not be used.

However, in a response to one of the articles most critical of incomplete factor scores (Glass and Maguire, 1966), Schweiker (1967) argued that the optimizing nature of the complete estimation weights would cause the com-

plete factor score's superior characteristics to disappear in favor of incomplete factor scores upon cross-validation. Wackwitz and Horn (1971) found this to be essentially true in an empirical examination of a fabricated data set with a simple structure; incomplete methods were found superior to complete methods on all three desirable internal characteristics upon double cross-validation.

It appears, therefore, that different types of factor score estimates have been compared on desirable internal characteristics, and that one type of complete factor score has been compared with full-rank data in replication predictive ability when they were used as independent variables in a regression equation. However, there seems to be no examination of the external characteristic of the predictive abilities of the several types of incomplete and complete factor scores in comparison with each other. Since there seem to be several benefits to the use of factor scores as predictors in multiple regression, the purpose of this study was to aid the researcher in selecting the most desirable type of factor score by comparing the predictive accuracy upon replication of regression equations created on several different kinds of complete and incomplete factor scores from a common factor analysis and varimax rotation of the independent variables.

METHOD

Factor score estimation methods. The five types of complete factor score estimates considered (see Harman, 1976, p. 368-384) were estimation by 1) regression, 2) theoretical model, 3) minimizing unique factors, 4) method 3 orthogonalized, and 5) ideal variables. The four incomplete methods were weighting by 1) the structure loading, 2) the loading squared, 3) positive or negative unit-weighting those variables with loadings larger in absolute

value than a specified salient value, and 4) unique unit-weighting with the restriction that a variable "loads" on only one factor, if at all.

Data generation. A data set was chosen from the literature for the empirical comparison of the performance of these different types of factor score estimates. It was thought that the most appropriate data might be some that had already been used in the literature to examine the internal qualities of the factor scores. Since the data used by Glass and Maguire (1966) seems relatively representative of what might be found in the literature, and because the article was among the most critical of incomplete factor scores, it was chosen as a representative and rigorous test case.

The necessary data from the Glass and Maguire study was the 13 variable full-rank intercorrelation matrix (Table 1) from which they compared the performance on internal characteristics of three of the four types of incomplete factor score methods mentioned previously, and the complete regression method. In addition, the amounts of multiple correlation on the full-rank data to be investigated were specified as R^2 = .30, .50, and .70; representing the range that might be found in the literature.

The procedure was to create a population of 1000 "subjects" by computer manifesting this predictor variable intercorrelation matrix and the desired multiple correlation. The correlation between the predictor variables and criterion in the population was specified as the same for all 13 variables and was chosen to yield the desired multiple correlation. This simple correlation which yields the desired multiple correlation can be shown to be equal to the square root of the sum of the elements of the inverse of the predictor variable intercorrelation matrix divided into the desired multiple correlation. Resulting simple correlations were .22, .29, and .34 respectively for the three multiple correlations. The method of creating a population with any desired correlation matrix has been described previously (Morris, 1975).

Comparison Procedure . The next step was to select a random sample and factor analyze the sample's 13 predictor variable intercorrelation matrix. A three factor (as in Glass and Maguire) principal axes solution was calculated in common factor space by iteration from initial squared multiple correlation communality estimates. This solution was then varimax rotated and factor score weights were created for the various methods. For the unit-weighting and unique unit-weighting methods, a salient loading was specified as one larger than .35, as in the Glass and Maguire study.

Next, a regression equation was created on each type of factor score predicting the criterion in the sample. Another sample was selected and parallel procedures were accomplished. At this point there was a regression equation from each of the two samples on each type of factor score. Each pair of regression equations on a factor score type was then double cross-validated by applying them to the alternate sample data. The index of accuracy for each of the various methods was the correlation of the predicted with the actual criterion score over both samples. Combined sample sizes of 42, 70, and 140 (representing subject to variable ratios of 3, 5, and 10 to 1) were investigated. For each of the nine combinations of multiple correlation and sample size, 100 iterations of this procedure yielded Fisher z transformed mean correlations taken as indicative of the methods' performance.

RESULTS AND CONCLUSIONS

While the sets of factor score weights generated by each of the various methods were different, the regression equations on factor scores resulting from the different methods were not all unique. By postmultiplying a factor score weight matrix by the vector of regression weights on the factor scores, a vector of weights is created which when multiplied by the full-rank data yields the regression prediction on the factor scores. In this way it was found

that regression on factor scores by the theoretical model, minimizing unique factors, and the orthogonalized minimizing unique factors yielded precisely the same regression equations. Further, the complete method of ideal variables yielded regression weights identical to the incomplete method of weighting by the structure loading. Thus for the purpose of regression prediction on factor scores there were only six unique methods.

Table 2 shows the mean correlations for the six factor score methods for each of the nine combinations of multiple correlation and sample size. In all cases the incomplete method of unique unit-weighting performed best and the most common complete method of estimation by regression performed poorest. In addition, each of the incomplete methods performed better than either of the two purely complete methods in all combinations. Since the ideal variable/structure loading method represented both complete and incomplete regression weights, it was not considered when comparing all complete with all incomplete methods.

Within a parameter combination, each of the 100 samples were operated on by each of the six factor score methods, thus constituting a repeated measures design with 100 analysis units, and six replicates. The Hotelling T^2 test for repeated measures (Morrison, 1967, p. 136) served as a test of overall equality as well as for certain *a posteriori* examinations. The F ratios resulting from the Hotelling T^2 (Table 3) suggested that there was an overall significant difference between means for each combination of multiple correlation and sample size.

Two further types of comparisons were of interest including a weighted contrast between the incomplete and complete methods, and a comparison of unique unit-weighting with the other five methods. The F ratios in Table 3 indicate that in every case, the incomplete factor scores performed significantly better as a group than the complete factor scores. In addition, for all combinations except for the smallest sample size with the two smallest multiple correlations, the unique unit-weighted factor score was significantly su-

erior to all the other five methods. It should be noted that, even in these ost unstable of sampling conditions of low multiple correlation and small mple size, the unique unit-weighted factor score still proved significantly uperior to the two complete methods (Table 3). It seems reasonable to expect hat given more samples, thus greater statistical power, the unique unit-weighted factor score would prove significantly superior in these two cases as well.

In addition, use of component analysis (ones as communalities) gave almost identical results to those from the common factor model. While there is always a possibility that results from a Monte Carlo study such as this are specific to the data set, this argument may also be applied to those studies cited which suggested from empirical research examples that incomplete factor scores not be used because of their inferior internal characteristics. Moreover, the reasonable ranges of relevant parameters one might encounter in the literature seem to be represented in this data.

It seems from the data examined that the simpler and more intuitively appealing incomplete factor score methods, especially the unique unit-weighted variety, are to be preferred over complete methods when the researcher wishes to derive a regression model capable of accurate prediction upon replication. Further, because multiple regression can be viewed as encompassing most parametric statistical procedures as special cases of its general model (Kerlinger and Pedhazur, 1973), it is proposed that these results may be applicable across several other statistical procedures.

REFERENCES

Burket, G. R. A study of reduced rank models for multiple prediction. Psychometric Monographs, 1964, No. 12.

Cattell, R. B. (Ed.) Handbook of multivariate experimental psychology. Chicago: Rand-McNally, 1966.

Glass, G. V., and Maguire, T. O. Abuses of factor scores. American Educational Research Journal, 1966, 4, 297-304.

Guertin, W. H., and Bailey, J. P. Introduction to modern factor analysis. Ann Arbor, Michigan: Edwards Brothers, 1970.

Halperin, S. The incorrect measurement of components. Educational and Psychological Measurement, 1976, 36, 347-353.

Harman, H. H. Modern factor analysis. Chicago: University of Chicago Press, 1976.

Horn, J. L. An empirical comparison of methods for estimating factor scores. Educational and Psychological Measurement, 1965, 25, 313-322.

Horn, J. L., and Miller, W. C. Evidence on problems in estimating common factor scores. Educational and Psychological Measurement, 1966, 26, 617-622.

Horst, P. Factor analysis of data matrices. New York: Holt, Rinehart, and Winston, 1965.

Kerlinger, F. N., and Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.

McDonald, R. P., and Burr, E. J. A comparison of four methods of constructing factor scores. Psychometrika, 1967, 32, 381-401.

Morris, J. D. A computer program to create a population with any desired centroid and covariance matrix. Educational and Psychological Measurement, 1975, 35, 707-710.

Morris, J. D. A comparison of regression prediction with data-level variables versus factor scores. Paper presented at the meeting of the American Educational Research Association, New York, April, 1977.

Morris, J. D., and Guertin, W. H. The superiority of factor scores as predictors. Journal of Experimental Education, 1977, 45, 41-44.

Morrison, D. F. Multivariate statistical methods. New York: McGraw-Hill, 1967.

Moseley, E. C., and Klett, C. J. An empirical comparison of factor scoring methods. Psychological Reports, 1964, 14, 179-184.

Schweiker, R. F. Factor scores aren't sacred: Comments on "Abuses of factor scores". American Educational Research Journal, 1967, 4, 168-170.

Tucker, L. R. Relations of factor score estimates to their use. Psychometrika, 1971, 36, 427-436.

Wackwitz, J. H., and Horn, J. L. On obtaining the best estimates of factor scores within an ideal simple structure. Multivariate Behavioral Research, 1971, 6, 389-408.

TABLE 1

Glass & Maguire's 13 Variable Intercorrelation* Matrix

	Variable												
	1	2	3	4	5	6	7	8	9	10	11	12	13
1	--	54	68	37	01	45	15	50	19	60	25	69	39
2	54	--	43	55	24	18	27	26	-08	24	37	38	00
3	68	43	--	49	14	64	24	72	45	55	21	84	55
4	37	55	49	--	42	30	36	31	01	08	44	36	-04
5	01	24	14	42	--	10	25	12	07	-21	13	07	-10
6	45	18	64	30	10	--	18	69	49	37	00	61	50
7	15	27	24	36	25	18	--	19	-06	-11	27	16	-12
8	50	26	72	31	12	69	19	--	47	44	06	67	55
9	19	-08	45	01	07	49	06	47	--	20	-29	45	51
10	60	24	55	08	-21	37	-11	44	20	--	11	64	49
11	25	37	21	44	13	00	27	06	-29	11	--	15	-21
12	69	38	84	36	07	61	16	67	45	64	15	--	57
13	39	00	55	-04	-10	50	-12	55	51	49	-21	57	--

* Decimal points omitted.

Table 2

Mean Double Cross-Validated Correlations for all Combinations of Multiple Correlation and Sample Size

Squared Multiple Correlation	Complete			Incomplete		
	Regression	Theoretical/ Mini. Unique Mini. Unique Orth.	Ideal/ Loading	Loading^2	Unit-Weighting	Unique Unit-Weighting
			N = 42			
.30	.0685	.1127	.1435	.1457	.1519	.1624
.50	.1709	.2422	.2885	.2889	.2838	.3086
.70	.2435	.3498	.4039	.4064	.4033	.4360
			N = 70			
.30	.1535	.1726	.2112	.2145	.2175	.2421
.50	.2615	.2958	.3513	.3513	.3581	.3871
.70	.3208	.3642	.4333	.4373	.4468	.4828
			N = 140			
.30	.2082	.2214	.2689	.2722	.2846	.3099
.50	.3116	.3326	.3928	.3962	.4109	.4492
.70	.3958	.4188	.4836	.4915	.5083	.5415

Note. Mean correlation from Fisher z transformation.

Table 3

$F(5,95)$ Values for Overall Tests of Equality of Methods and A Posteriori Contrasts for All Combinations of Multiple Correlation and Sample Size

Squared Multiple Correlation	Overall	Contrasts with Unique Unit-Weighting					Complete vs. Incomplete Contrast
		Regression	Theoretical/ Mini. Unique/ Mini. Unique Orth.	Ideal/ Loading	Loading	Unit-Weighting	
			N = 42				
.30	12.3**	9.5**	6.9**	1.6	1.9	.4	11.7**
.50	25.1**	15.6**	8.6**	1.6	2.2	2.5*	16.5**
.70	35.4**	22.9**	24.9**	6.3**	7.8**	5.6**	27.7**
			N = 70				
.30	37.7**	34.3**	26.4**	8.6**	10.7**	5.4**	33.4**
.50	127.3**	84.8**	50.1**	11.6**	20.5**	8.9**	87.0**
.70	123.7**	101.9**	91.5**	35.3**	32.3**	17.1**	120.5**
			N = 140				
.30	119.8**	82.1**	68.3**	27.0**	32.6**	14.2**	97.3**
.50	312.1**	247.9**	198.4**	78.4**	80.7**	37.2**	276.9**
.70	367.8**	229.3**	163.6**	62.4**	62.0**	17.3**	208.3**

* $p < .05$

** $p < .01$